

## ETSO Workshop:

# EASA AI Roadmap: Towards AI trustworthiness...

**Guillaume Soudain**  
EASA Programme Manager – Artificial Intelligence

21<sup>st</sup> September 2022



# EASA AI Roadmap – Towards AI trustworthiness

- Impact on all aviation domains
- Common issue: level of trust in AI performance
- « AI trustworthiness » concept is the key

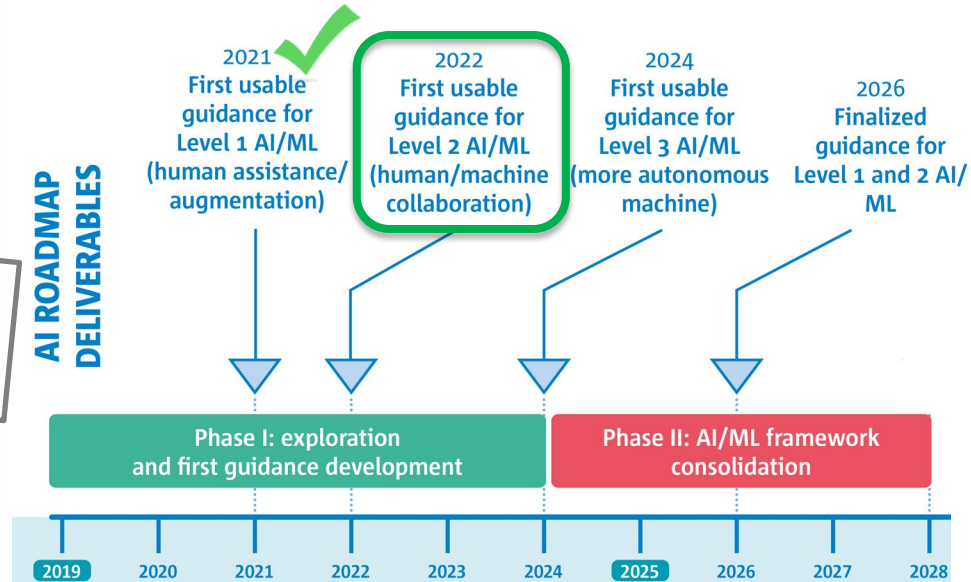


**EASA Concept Paper:**  
First usable guidance for Level 1  
machine learning applications

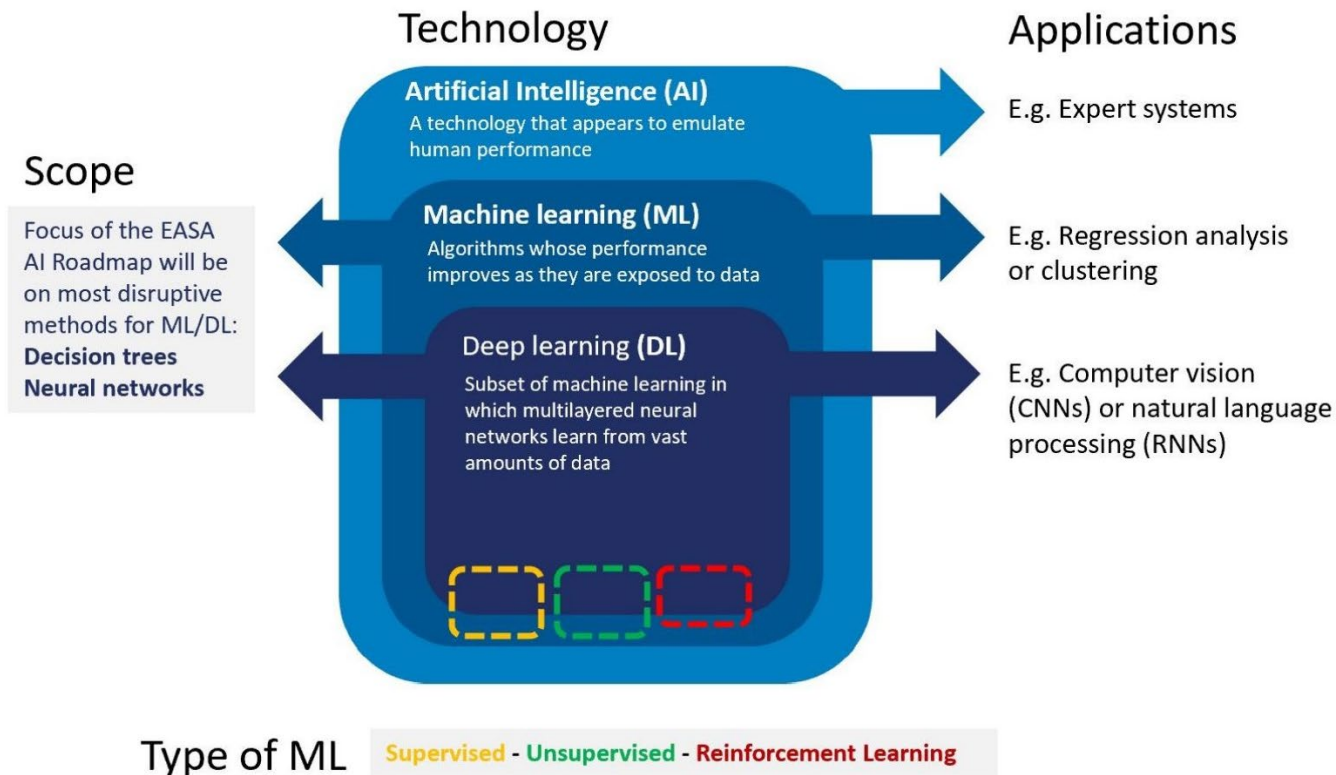
A deliverable of the EASA AI Roadmap

**Published on 20<sup>th</sup> December 2021**  
<https://easa.europa.eu/ai>

April 2021  
Proposed Issue 01



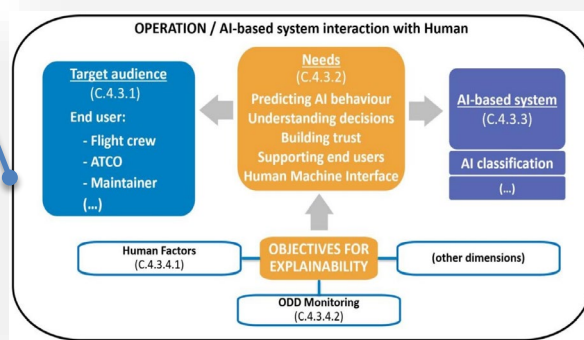
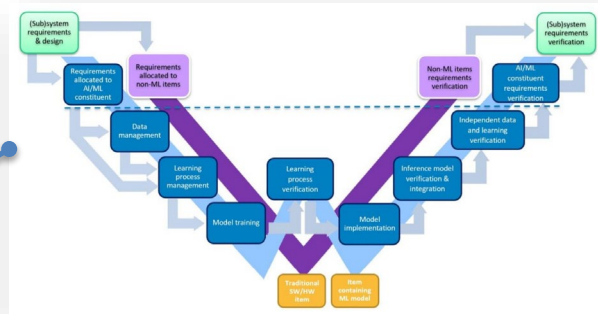
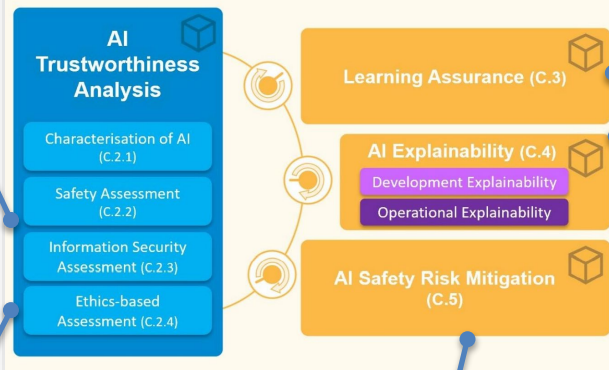
# EASA AI Roadmap – Definitions and Scope




# First usable guidance - Overview



## EASA Trustworthy AI building blocks



# AI classification scheme



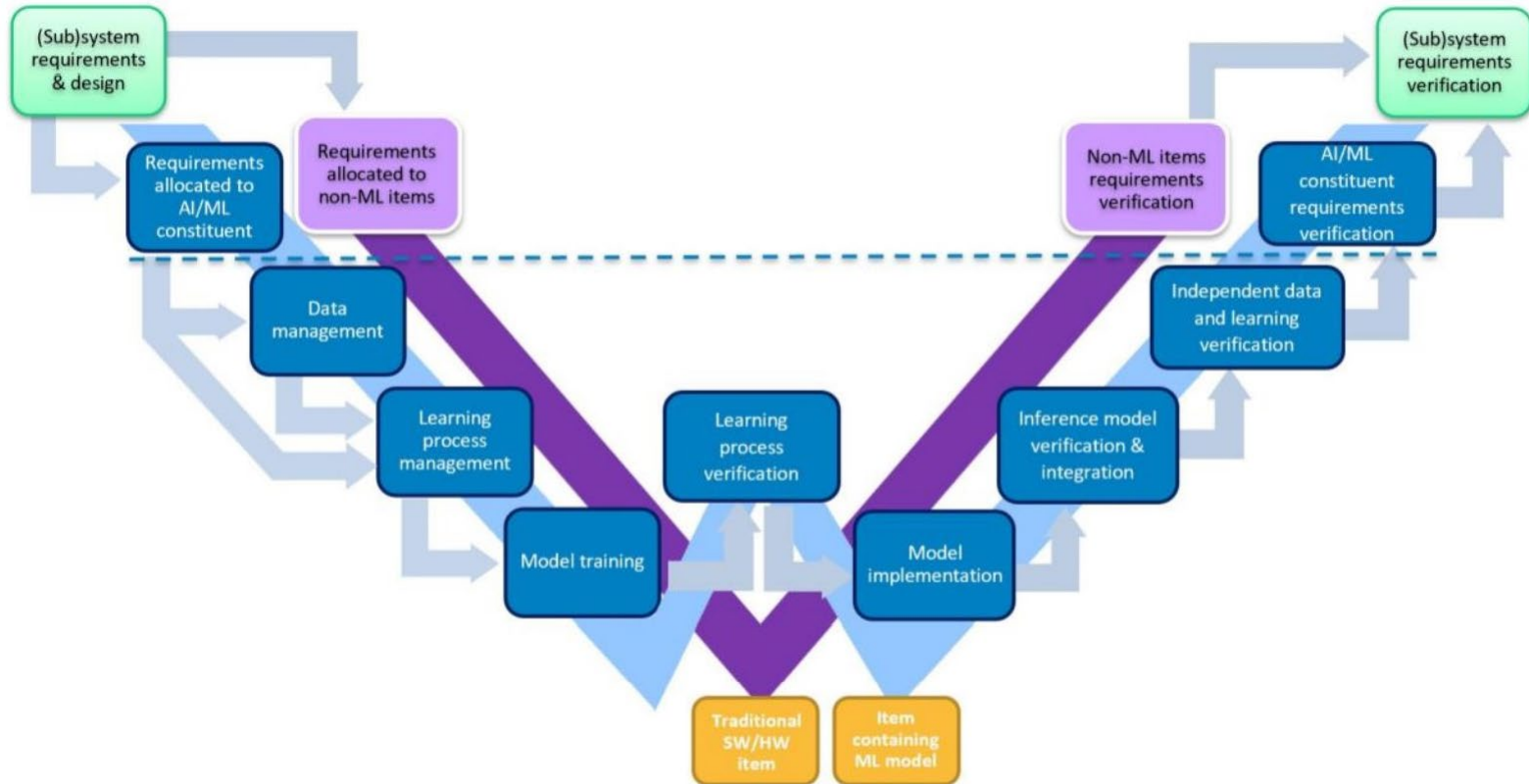
Increasing Automation

| EASA AI Roadmap<br>AI Level          | Function allocated to the system to contribute to the<br>high-level task |
|--------------------------------------|--|
| Level 1A<br>Human<br>augmentation    | Automation support to information acquisition                            |
|                                      | Automation support to information analysis                               |
| Level 1B<br>Human assistance         | Automation support to decision-making                                    |
| Level 2<br>Human-AI<br>collaboration | Overseen and overridable automatic decision-making                       |
|                                      | Overseen and overridable automatic action implementation                 |
| Level 3A<br>More autonomous AI       | Overridable automatic decision-making                                    |
|                                      | Overridable automatic action implementation                              |
| Level 3B<br>Fully autonomous AI      | Non-overridable automatic decision-making                                |
|                                      | Non-overridable automatic action implementation                          |

Table 1 — EASA AI typology and definitions

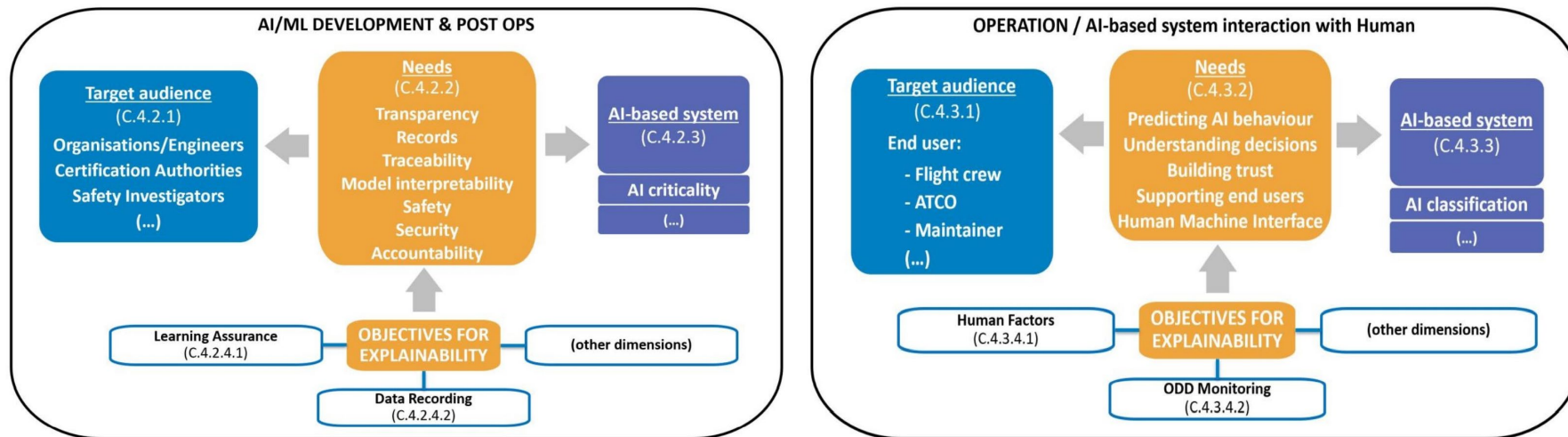


# Learning Assurance: W-shaped process



# AI Explainability – Overview

**AI explainability:** Capability to provide the human with understandable, reliable, and relevant information with the appropriate level of details and with appropriate timing on how an AI/ML application is coming to its results.



# 2022+ Main challenges for Level 2 AI/ML guidance

## 1. Anticipate means of compliance for Learning Assurance objectives on ML Model guarantees (generalization and robustness)

→ Exploit the Horizon Europe Research project MLEAP on 'Machine Learning applications approval'

## 2. Operational explainability & human centric aspects of AI

→ Foster confidence in the system by developing specific HF guidance

## 3. Ethics-based assessment – social & societal aspects

→ Evaluate and refine guidance based on use cases

Partnering on use cases is a key driver!



# Exercising the AI Guidance with use cases



AI/ML Guidance



Use Cases



Standardisation  
WG-114/G-34



# Thank you for your attention!

[ai@easa.europa.eu](mailto:ai@easa.europa.eu)

**TRUST**

[easa.europa.eu/connect](https://easa.europa.eu/connect)



**Your safety is our mission.**

An Agency of the European Union 